

8 Assessing Language Proficiency and Intercultural Development in the Overseas Immersion Context

Dan E. Davidson, Nadra Garas and Maria D. Lekic

Introduction

Among intercultural competencies, language plays a critical role, central to the perception and interpretation of external reality, the construction and transfer of meaning and the projections of self. Contemporary information technologies have done much to minimize geographical and temporal barriers to global communication, but native language(s), local culture and regional context strongly influence all modes and modalities of intercultural communication, particularly where the establishment or maintenance of mutual understanding, personal trust and professional collaboration are concerned. For that reason, overseas language immersion is an obligatory component of all Language Flagship programs, as well as of many other international programs today. Second language (L2) gain within the context of well-designed immersion study is extensively documented in the literature, as is the need for the overseas sojourner to comprehend and engage effectively with the host culture (Davidson, 2015).

Alumni of successful overseas immersion learning programs such as the Flagship are able to consciously adjust cognitive perspectives and adapt to different cultural environments, utilizing appropriate sociolinguistic and intercultural communication strategies. These abilities are what is meant by *intercultural competence* (IC), as reflected in Deardorff's (2004) widely recognized 'consensus definition'. IC comprises both internal (cognitive and attitudinal) and interactional components for an adaptable and nuanced ethno-relative capacity to interpret contexts and function effectively across cultural divides (Deardorff, 2006; King & Baxter Magolda, 2005). The interrelationship

of second language acquisition (SLA) and intercultural development in the study abroad (SA) context, while often acknowledged as relevant, has remained largely unspecified. One recent study, for example, found that pre-program intercultural development scores appear to predict ultimate L2 proficiency outcomes (Baker-Smemoe *et al.*, 2014), while other researchers have observed parallel development of IC and L2 proficiency (Paige *et al.*, 2004; Watson *et al.*, 2013). Most existing research on IC and SA, however, has been limited to the examination of IC measurements of consecutive groups of SA populations who are at roughly the same stage in their learning careers and, typically, in connection with a first academic sojourn abroad, whether for a summer, semester or academic year. It is difficult in most cases to account for the differential effects of language proficiency, intercultural orientation and academic background on the overall outcome of the immersion experience.

The Intercultural Development Continuum (IDC) (Hammer, 2007, 2012; Hammer & Bennett, 1998, 2002), based on the intercultural sensitivity model first elaborated by Bennett (1986, 1993, 2004), characterizes orientations to cultural difference. Hammer (2012; Appendix 8.1) adapted the IDC to five measureable orientation ranges assessed by the Intercultural Development Inventory (IDI). The score produced by the IDI places each individual or group along a continuum ranging from monocultural perspectives to more intercultural/global mindsets: Denial, Polarization, Minimization, Acceptance and Adaptation.

A number of studies have used the IDI to examine the impact of SA programs. Vande Berg *et al.* (2004) employed the IDI to compare gains in intercultural development by SA students with gains of those students who had remained on the home campus over the same period, observing that the SA group generally exhibited more growth in intercultural development as well as some competence in an L2 in comparison with those who did not go abroad. Paige *et al.* (2004) studied the impact of the SA experience and curriculum intervention on students' intercultural development, SLA and employment of learning strategies related to language and culture, noting overall gains in intercultural sensitivity. Engle and Engle (2004) identified higher intercultural gains among SA students whose programs also included a cultural mentoring component as well as direct contact with the host culture.

Paige *et al.* (2004) and Engle and Engle (2004) examined SA language learning programs and intercultural sensitivity, concluding that those students who spent a year overseas showed gains in intercultural sensitivity, while those engaging in shorter-term programs were less likely to exhibit gains in IDI. A study by Medina-Lopez-Portillo (2004) also found little support for significant gains over short programs (seven weeks to one semester). However, Anderson *et al.* (2006) found that

students who participated in two- and four-week SA programs showed small gains in intercultural sensitivity, compared with no gains for those students who spent that same time enrolled in courses at their home universities. Patterson (2006) examined the impact of short-term SA programs (two- or four-week programs) and concluded that those students who engaged in SA showed small improvements in intercultural effectiveness compared with those who spent that time in the classroom. Rexeisen *et al.* (2008) concluded that SA programs have a positive short-term effect on intercultural development, but noted that assessment of the long-term impact required further investigation. Anderson and Lawton (2011) argue that SA exerted a positive change on the intercultural development of students, as supported by the results of the pre-post administration of IDI to undergraduate students who spent a one-semester SA program in London.

Paige *et al.* (2004) and Vande Berg (2009) conclude that students studying abroad show gains in intercultural learning, using IDI scores to compare the impact of various curricular strategies adopted by SA programs on their students' IC. The Georgetown Consortium multi-year study showed that students in SA made small gains, but that these were tempered by the duration or length of the SA program and the extent to which the program offered guided learning. In some cases, the study revealed no significant difference between SA students and those who stayed at their home universities (Vande Berg *et al.*, 2009). Similarly, Pedersen (2010) concludes that guided learning is central to intercultural development and learning experience and that having students spend time overseas is not sufficient to foster effective global citizenship.

Among studies linking SA and intercultural development, Salisbury's (2011) large-scale ($N=1593$) longitudinal examination is of particular note. Drawing on data collected in 2006 for the Wabash Study of Liberal Arts Education, Salisbury (2011: 92) establishes, under rigorous analytic conditions, a statistically significant positive effect for SA on IC, 'an effect that appears to be general, rather than conditional'. While establishing that SA influences students' diversity of contact, the study was unable to establish statistically significant effects on participants' relativistic appreciation of cultural differences or comfort with diversity (Salisbury, 2011). The Salisbury study does not control for duration of immersion, study designs or engagement by participants in language or area studies.

Watson *et al.* (2013) assess change in intercultural development (using the IDI), language proficiency (oral proficiency interview [OPI]) and area knowledge (ARK/GRANT)¹ in a large-scale ($N=498$) longitudinal examination of third- and fourth-year-level US Military Academy cadets who studied one of seven different languages for a semester in 14 overseas study locations. The study reports multimodal proficiency gains at the unit

level (e.g. American Council on the Teaching of Foreign Languages [ACTFL] Intermediate Low to Intermediate Mid) and 49% at the threshold level (e.g. Intermediate to Advanced) along with modest IDI growth across groups for semester-long cohorts placed for language study in China, Eastern Europe, Western Europe, Latin America and the Middle East (Watson *et al.*, 2013). The study also reports mean IDI gains in developmental orientation ranging from 2.5 to 5.1 points (87.2–92.3) over the four consecutive years for which data were collected. These changes are consistent with others reported in the literature for US undergraduate SA (Hammer, 2012; Vande Berg, 2009).

Baker-Smemoe *et al.* (2014) present data for approximately 100 summer and semester-long SA university students with initial measured proficiency levels largely in the ACTFL Intermediate range in six world languages. In comparing those whose proficiency increased over the course of the SA program with those who did not register gains, the authors report a positive correlation between pre-program IDI developmental orientation scores and ultimate gain (or no gain) across all SA groups under study. Among those variables assessed, pre-program IDI was shown to have the largest predictive value for language gain.

Research Questions

The construct underlying the assessment of intercultural competencies comprises a mix of traits related to the learner's background knowledge, critical thinking abilities, attitudes, executive functioning and interpersonal skills. The present study hypothesizes that language proficiency both influences and is influenced by the learner's level of intercultural development. Building on existing research on language gain and taking into consideration the range of evidence that has appeared regarding learning in the overseas immersion context, the present study poses the following research questions:

- (1) How does OPI change differ across the immersion programs focused on early, mid and advanced levels of training and what, if any, relationships obtain between different levels of language proficiency and measured levels of intercultural development, as reflected by the IDI?
- (2) To what extent do pre-program IDI scores predict ultimate L2 attainment of SA participants?
- (3) What is the effect of lower, mid-range and advanced L2 proficiency on change in IDI during SA?

While the 'language barrier' is popularly cited as the most challenging factor to successful interactions across cultures, the present study seeks to better understand the role of proficiency in an L2 and the learner's

own intercultural development of language as a potential barrier and as a potential mediator of interculturality.

Methodology

The study will first report and provide statistical comparisons of pre- and post-program OPI and IDI scores of students of various foreign languages who have undertaken formal language and cultural immersion training in countries where those languages are spoken. Change (positive and negative) in pre- to post-program OPI and IDI ratings will be noted and significant relationships highlighted. Significant sub-score effects, where observed, will also be reported.

The OPI is widely used in the United States today to measure speaking proficiency, based either on the government Interagency Language Roundtable (ILR) scale or on the ACTFL scale (Appendix A). The IDI is an online instrument which assesses intercultural development orientation (DO), perceived intercultural orientation (PO) and the orientation gap (OG: the gap between DO and PO) along a scale ranging from monocultural mindsets (Denial) up to and including a level of intercultural awareness and sensitivity sufficient for functional integration into the host culture (Adaptation; Appendix 8.1). The 50-item IDI generates a profile and a set of numerical scores and sub-scores for each test taker, placing the test taker on the IDI continuum. Movement along the continuum in either direction is possible. The IDI is available commercially and is used widely in government, industry and higher education today.

Subject Groups

Participants in the present study ($N=305$) consist of US high school and university students selected for participation in 2013 and 2014 for four federally supported overseas language immersion programs: Flagship capstone overseas program students ($N=44$), Russian Language and Areas Studies (RLASP) semester-long participants ($N=104$), National Security Language Initiative for Youth (NSLI-Y) ($N=132$) and outbound Kennedy Lugar Youth Exchange and Study Abroad program (YES) ($N=25$). YES and NSLI-Y students are comparable in age, selection, background and duration of overseas immersion (one year). The first three groups are engaged in formal language study. YES students are not engaged primarily in language study, but enjoy similar kinds of in-country support provided by English-speaking hosts. The NSLI-Y participants took part in year-long programs in China, India, Korea, Moldova, Morocco and Taiwan. The RLASP students participated in semester-long programs in various urban university locations in Russia; the Flagship capstone students participated in year-long programs in China, Kazakhstan, Morocco and Turkey and the YES

students participated in year-long programs in Bosnia-Herzegovina, Ghana, Indonesia, Jordan, Macedonia, Morocco, Oman, Turkey and South Africa, residing with English-speaking families of recently returned inbound YES exchange students to the United States. Full demographic information is provided in Table 8.1.

Table 8.1 Demographic information

<i>All programs</i>	<i>L-2 programs*</i>	<i>NSLI-Y</i>	<i>RLASP</i>	<i>All Flagship programs</i>		<i>YES program</i>
				<i>(ROF only)</i>		
N = 305	280	132	104	44	(19)	25
Age						
17 and under	21 (7.5%)	21 (15.9%)	–	–	–	21 (84%)
18–21	202 (72.1%)	111 (84.1%)	86 (82.7%)	5 (11.4%)	1 (5.3%)	4 (16%)
22–30	56 (20%)	–	18 (17.3%)	38 (86.4%)	17 (89.5%)	–
31–40	1 (0.4%)	–	–	1 (2.3%)	1 (5.3%)	–
Gender						
Male	117 (41.8%)	53 (40.2%)	40 (38.5%)	24 (54.5%)	11 (57.9%)	4 (16%)
Female	163 (58.2%)	79 (59.8%)	64 (61.5%)	20 (45.5%)	8 (42.1%)	1 (84%)

* L2 programs (N=280) are Flagship. NSLI-Y and RLASP. Russian Overseas Flagship (ROF) is a subset of Flagship. The YES outbound program (N=25) does not require previous language study.

The limitations of the participant samples should be clearly noted from the outset. The participants cannot be said to represent the US population more generally, or US high school or university students for that matter, due to the impossibility within the context of four competitive federal programs of controlling for selection effects or producing a randomized study of treatment or control groups. To the extent, however, that all four cohorts were selected for participation using widely accepted standardized criteria that did *not* consider IDI scores or consideration of the student's financial situation, and that participants represent a very broad range of public and private institutions, they may be regarded as typical of US students who currently study overseas at some point in their high school or university careers. In that sense, the results of this study should be useful for academic planning, policy and SLA research purposes.

Analysis²

Pre- and post-program OPI results

Figures 8.1 through 8.3 present pre- and post-program OPI results of the three programs in question. NSLI-Y accepts some students with no prior study of the language in question, whereas RLASP and Flagship have specific qualifying levels for entering participants of ILR 1 and ILR 2 proficiency, respectively.

Within the NSLI-Y cohort for 2014, 61% of participants completed the program at ACTFL Advanced (ILR 2) or higher, while an additional 23% scored at Intermediate High (IH), which is at or near the threshold of ILR 2 (Figure 8.1).

RLASP programs are of one-semester duration. It is noteworthy that 50% of the 2014 RLASP cohort completed the program at ILR 2, with an additional 24% of the program at IH, at or near the threshold of ILR 2 (Figure 8.2).

Language Flagship overseas programs accept students from US domestic Flagship programs who have attained ILR 2 proficiency in two modalities and who have devoted a minimum of eight weeks of prior formal academic study in the target country at the time of application. It should be noted, however, that within the present cohort, 10% of the total group were rated ACTFL IH rather than Advanced on the pre-program OPI. For the 2014 cohort represented in the study, 70% of students achieved ILR 3 (or higher) in speaking on the post-program OPI, while 23% were rated ACTFL AH, i.e. at or near the ILR 3 threshold (Figure 8.3).

The Russian Overseas Flagship (ROF) in 2014–2015 was shifted from St. Petersburg State University (Russia) to Al-Farabi Kazakh National

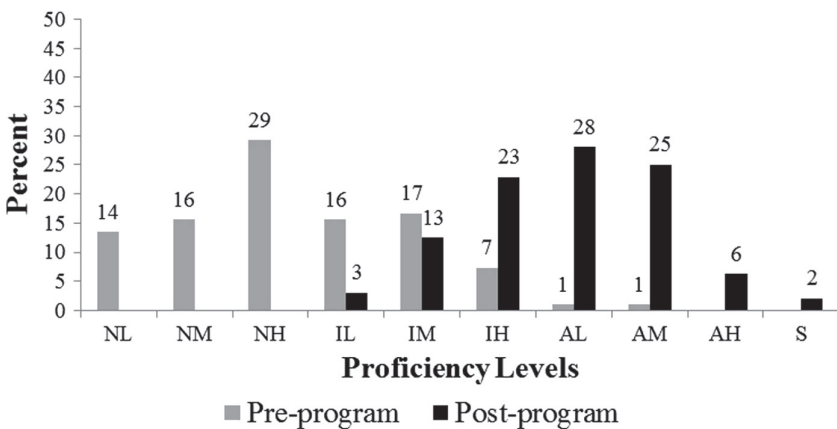


Figure 8.1 NSLI-Y programs pre- and post-program OPI levels ($N=96$)

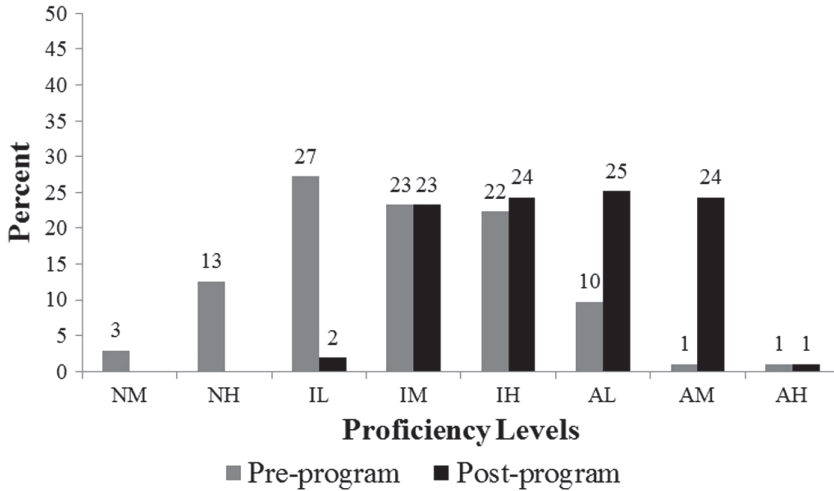


Figure 8.2 RLASP program pre- and post-program OPI levels (N=103)

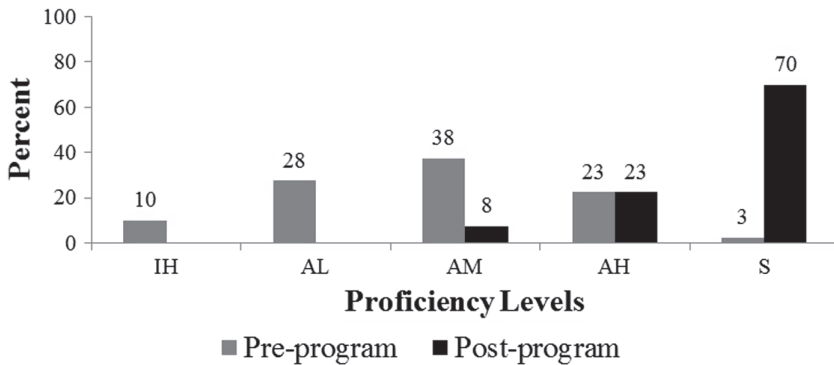


Figure 8.3 All Overseas Flagship programs pre- and post-program OPI levels (N=40)

University in Almaty (Kazakhstan). While still very much a part of the Russophone world, Kazakhstan is a bilingual nation in the heart of Central Asia, creating a tri-cultural experience for the 2014–2015 ROF cohort. For that reason, the ROF results are reviewed separately in Figure 8.4.

Analyzing OPI change across the three programs

The effects of the three language immersion programs on OPI growth were examined with a mixed linear model. The fit statistics for a series of related models varying in complexity are given in Table 8.2.

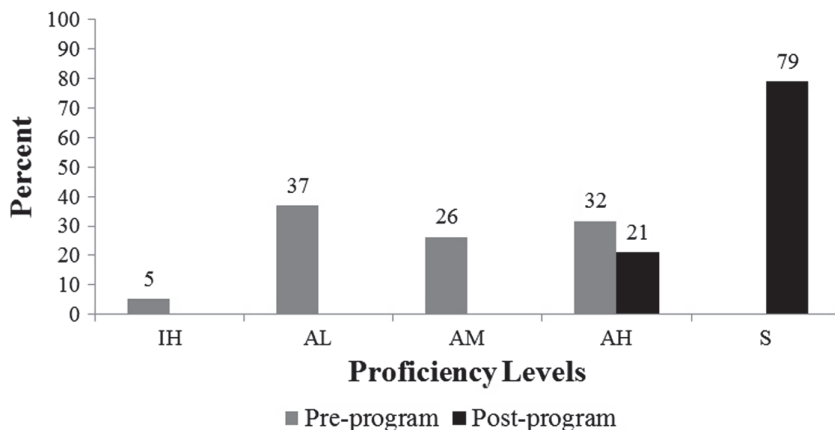


Figure 8.4 Russian Overseas Flagship Program pre- and post-program OPI levels ($N=19$)

The model in Table 8.2 is the best-fitting linear model of pre- and post-program OPIs. It indicates that the fixed-effect interactions between growth (f1) and program type are significant. While OPI change differences exist across the three programs, the average OPI change is approximately the same across all three programs. Language Flagship students attain comparable growth as lower- and middle-range-proficiency learners, but they produce these gains in the upper ranges of functional proficiency (ILR 3). The NSLI-Y group starts with the lowest performance (of $7.80-4.46=3.34$) and demonstrates the largest improvement (of $1.83+1.63=3.46$). These interaction effects are several times larger than their standard errors and thus highly significant.³

Table 8.2 A mixed linear model of OPI growth

<i>Model</i>	<i>AIC</i>	<i>DIC</i>	<i>Deviance</i>	<i>Resid (SD)</i>
1. Between Subjects Only	2262.7	2251.3	2254.0	2.00
2. Between Subjects+Growth	1958.4	1939.3	1944.8	1.08
3. Between Subjects+Program	2095.6	2073.7	2079.6	1.83
4. Between Subjects+Program+Growth	1762.6	1731.2	1740.9	1.09
5. Between Subjects+Program*Growth	1672.8	1630.9	1643.9	0.89

Pre- and post-program IDI scores

Table 8.3 presents the results of IDI data collected before and after participation by students in the three language-based immersion programs, as well as for the YES program. The DO score is the basal score produced by the IDI, against which a PO score (based on the same scale) is juxtaposed. The 'gap' measurement, the difference between DO

and PO, is also a part of the IDI profile and is a significant finding in the present study.

Table 8.3 IDI pre- and post-mean developmental orientation scores

<i>Program name</i>	<i>Mean pre</i>	<i>SD pre</i>	<i>Mean post</i>	<i>SD post</i>	<i>Mean difference</i>
ROF program only (<i>N</i> =19)	97.11	16.34	102.26	14.11	5.15
OPI ILR 3 and above	97.92	16.47	103.05	14.13	5.14
All Flagship programs (<i>N</i> =42)	97.95	15.41	99.08	13.13	1.13
RLASP program (<i>N</i> =103)	95.53	12.78	97.84	14.91	2.31
NSLI-Y program (<i>N</i> =131)	92.66	13.73	93.94	13.85	1.28
YES program (<i>N</i> =25)	95.32	12.02	95.48	14.32	0.16

Group-level DO scores increased in every group taking part in the overseas immersion study, with the smallest mean difference between pre- and post-program score of 0.16 registered for the non-language YES cohort and the largest mean difference of 5.14 reported for the Russian Overseas Flagship Program. It is worth noting that the YES group entered the program with mean IDI scores of 95.32, a relatively high mean pre-program level in comparison with others reported in the literature for SA students in general. DO scores show large within-group-by-occasion standard deviations of 14–16, and together with a small group size (19) for ROF make several effects more difficult to detect. Analysis of variance (ANOVA) comparisons did not indicate any group differences before ($F=2.405$, $df=[2, 277]$, $p=0.09$) or after ($F=2.88$, $df=[2, 277]$, $p=0.06$) program participation. RLASP shows a detectable change from 96 to 98 (difference=2.31, $t=1.99$, $df=103$, $p=0.048$). The data also indicate an increase of the DO scores among the ROF students, but because of the small group size ($N=19$), additional data will have to be collected to affirm this trend. For the current ROF data, (pre=96, post=101, diff=4.83; $t=1.975$, $df=18$, $p=0.06$).

PO scores are scaled similarly as DO. Pre-program group means vary between 122 and 124; the post-program means vary between 123 and 124. The program groups do not show any detectable mean differences either before ($F=1.81$, $df=[2, 277]$, $p=0.17$) or after ($F=1.203$, $df=[2, 277]$, $p=0.30$) program participation. Across all programs, the PO scores show an increase from 123 to 124 (diff=1.08, $t=3.599$, $df=279$, $p<0.001$). Within individual programs, NSLI-Y shows a significant increase from 122 to 123 (diff=1.049, $t=240$, $df=131$, $p=0.018$) and RLASP from 123 to 124 (diff=1.357, $t=2.939$, $df=103$, $p=0.004$).

OG is defined as the difference between PO and developmental scores. Means and standard deviations of these data are different from the DO and PO scores from which they are derived – the group-by-occasion means range

between 26 and 29, while the associated standard deviations range between 8 and 10. Pre-program means range between 26 and 29 and there is no systematic difference between the groups ($F=2.829$, $df=[2, 277]$, $p=0.06$). The largest OG pre-program mean (29) is found for the NSLI-Y group. Post-program OG means range between 26 and 29, and significant differences are apparent ($F=4.002$, $df=[2, 277]$, $p=0.019$). The RLASP and Flagship groups have the smallest post-program means of 26, while the NSLI-Y program shows an overall post-program mean of 29. The ROF program, taken separately, shows a drop of 3.26, from 27 to 24. The significance of this difference is indicated by $t=-2.166$, $df=18$, $p=0.04$.

The IDI profile also includes sub-scores for specific measureable traits, two of which revealed significant results in the current study (Table 8.4).

Pre-program *cognitive frame shifting* (Cog) scores range from 3.6 to 3.8 but are not statistically different ($F=2.593$, $df=[2, 277]$, $p=0.08$). The post-program Cog scores, however, do show some differences, with the YES score of 4.32, the NSLI-Y score (3.95) the highest and the RLASP scores the lowest (3.70) ($F=5.379$, $df=[2, 277]$, $p=0.005$). Although this difference is in the decimal figures (0.37), it reflects one quarter of the within-group-by-occasion standard deviation. There is also a significant Cog score increase across all data ($diff=0.17$, $t=4.293$, $df=279$, $p=0.000$). This increase can be detected in both the NSLI-Y ($diff=0.20$, $t=3.502$, $df=131$, $p<0.001$) and RLASP ($diff=0.13$, $t=2.014$, $df=103$, $p=0.047$) programs, but currently not with Flagship.

Table 8.4 Paired sample *t*-tests for YES program ($N=24$)

Variable	Mean pre	SD pre	Mean post	SD post	Mean difference	<i>t</i>	<i>p</i>
Cognitive frame shifting (Cog)	3.72	0.68	4.32	0.56	0.60	3.928	0.001
Behavioral code shifting (Beh)	3.96	0.61	4.44	0.65	0.48	3.361	0.003

Pre-program group mean scores of *behavioral code shifting* (Beh) show systematic differences ($F=13.28$, $df=[2, 277]$, $p<0.001$). RLASP shows the smallest pre-program mean (3.57), Flagship the largest mean (4.02) and the means of YES and NSLI-Y fall in between (3.96 and 3.95, respectively). The post-program group means also differ significantly ($F=9.861$, $df=[2, 277]$, $p<0.001$). RLASP again has the smallest mean (3.77), YES and NSLI-Y show the largest post-program mean (4.44 and 4.11, respectively) and the Flagship programs fall in between (3.97).

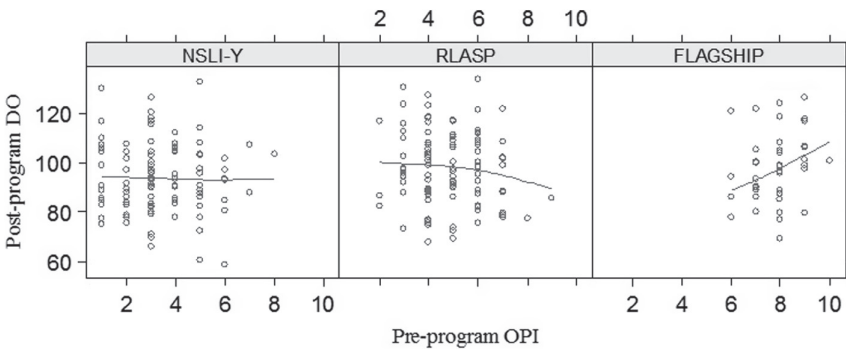
Beh scores increase significantly across all programs ($diff=0.15$, $t=4.710$, $df=279$, $p=0.000$). The increase is detectable for YES and NSLI-Y students ($diff=0.19$, $t=4.149$, $df=131$, $p=0.000$) and RLASP students ($diff=0.20$, $t=3.711$, $df=103$, $p<0.001$) but not among the Flagship students (Table 8.5).

Table 8.5 Paired sample *t*-tests for NSLI-Y program (*N*=131)

<i>Variable</i>	<i>Mean pre</i>	<i>SD pre</i>	<i>Mean post</i>	<i>SD post</i>	<i>Mean difference</i>	<i>t</i>	<i>p</i>
Acceptance	3.839	(0.717)	4.218	0.654	0.379	4.688	0.000
Adaptation	3.836	(0.530)	4.056	0.540	0.220	3.257	0.002
Cognitive frame shifting	3.701	(0.703)	3.947	0.645	0.246	2.658	0.010
Behavioral code shifting	3.945	(0.546)	4.142	0.551	0.197	2.899	0.005

Modeling the interrelationship of OPI and IDI

The trellis plots in Figures 8.5 through 8.7 represent regressions performed on NSLI-Y, RLASP and Flagship pre- and post-program values plotted against OPI or IDI by program type. (The YES program is not included in this portion of the analysis.) Figure 8.5 models the relative power of pre-program OPI ratings to predict post-program developmental orientation score. Numbers corresponding to proficiency levels (0 to 10, where 10=Superior) form the horizontal axis, while the vertical axis reflects the post-DO IDI score range. The NSLI-Y OPI values are centered primarily in the two left quadrants of the box graph, representing pre-program proficiency levels in the 0 to ACTFL Intermediate range, while the undergraduate RLASP scores are distributed along both sides of the central axis, representing pre-program OPIs in the ACTFL Intermediate range, annotated here as 3–7 on the horizontal axis. The Flagship box with scores concentrated in the two right-hand quadrants represents pre-program proficiencies in the ACTFL Advanced range (6–10). As is readily evident

**Figure 8.5** Post-program developmental orientation versus pre-program oral proficiency

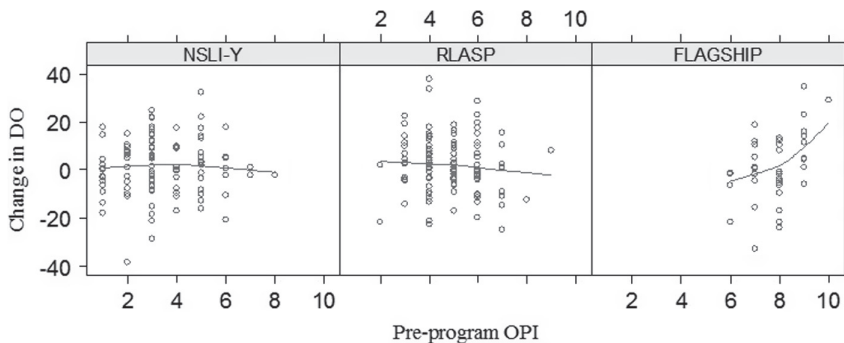


Figure 8.6 Change in developmental orientation versus pre-program oral proficiency from the slopes of the three regression plots, NSLI-Y and RLASP pre-program OPI scores do not correlate with DO growth, whereas Flagship pre-program OPI scores are correlated with DO gain.

An IDI-DO ‘change’ variable was then created to assist in further analyzing developmental trends in IDI performances across programs. The vertical axis in Figure 8.6 represents change (either positive or negative), while the horizontal axis marks pre-program OPI, as in Figure 8.5. The positive relationship between OPI and DO gain is evident in the upper right-hand quadrant of the Flagship panel in Figure 8.6.

Figure 8.7 compares pre-program DO with post-program OPI outcomes for the 280 language-immersion participants at each of the three levels of study represented by NSLI-Y, RLASP and Flagship. As is evident from all three panels, no relationship between pre-DO and post-program OPI was detectable. A further negative finding was produced when pre-DO and a post-program OPI gain variable was tested. The finding proposed recently by Baker-Smemoe *et al.* (2014) could, therefore, not be replicated in the present study.

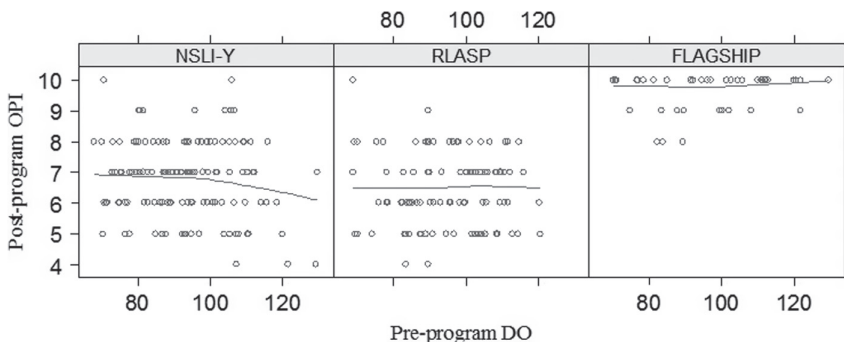


Figure 8.7 Post-program oral proficiency versus pre-program developmental orientation

Discussion

The powerful impact of structured immersion training on L2 gain has been reported in earlier studies (Davidson, 2015; Davidson & Lekic, 2012). These data and the accompanying analyses of OPI gain across levels of study and different languages are entirely new and reported here for the first time. They are consistent with earlier reports on these federal programs.

In addition to providing a comparative analysis of L2 gains in early, middle and upper ranges of proficiency of several critical languages, the present analysis also focused on the relationship between L2 gain at each level and the development of intercultural competencies within an overseas structured immersion context. In order to do this, it has made use of two widely recognized assessment tools, the ACTFL OPI and the IDI, and collected pre- and post-program data ($N=305$) using both instruments for groups of learners of different languages from a broad range of US institutions, who are at the early, middle or later stages of formal study of the language while engaged in year-long or semester-long structured L2 immersion programs overseas. In addition, we have reported on a further cohort of US SA students (students on the YES program), whose overseas study time was not focused on the formal study of language. The overall developmental orientation levels presented here, as reflected by IDI scores along the continuum, fall generally within the Minimization orientation (Appendix 8.1). As noted, mean IDI scores showed a positive change, from 92.11 to 103.05 across all cohorts, which indicates growth both within a single (Minimization) range, or, in some cases, post-program placement at the Acceptance orientation. The negative change in the OG (from 29.58 to 23.75) indicates a clear movement toward narrowing the gap between participants' perceived sense of intercultural sensitivity and their overall sensitivity as measured by the IDI. The heightened and more accurate level of intercultural self-awareness reported here is reminiscent of the increasingly accurate levels of oral proficiency self-assessment that has been observed among L2 speakers at or near ILR 3 and above (Freed *et al.*, 2004).

These results point to a significant change for all participants in the levels of Acceptance and Adaptation measured by the IDI. As noted, results of the sample *t*-test indicate a statistically significant change in the direction of growth for the IDI subscales of Acceptance and Adaptation, within the clusters of Cog and Beh. Moreover, mean scores on the Acceptance subscore orientation among NSLI-Y participants changed from a pre-test score of 3.84 to a post-test score of 4.22 ($p \leq 0.05$). Within the Acceptance orientation, individuals recognize and appreciate not only similarities but, more critically, differences within cultures.

Overall, participants demonstrate increased cultural sensitivity and competence at the end of their SA program. All four programs facilitate

and guide cultural learning through pedagogies and support structures that provide teacher and peer mentoring, monitored homestays, direct instruction and regular periods of self-reflection and self-assessment, interventions that are now widely recognized as best practices within the SA community. For example, students in one of the cohorts were asked to evaluate their own efforts to engage with the local population and context (on a scale of 1 to 10 where 10 was the highest rating) toward the end of their sojourn. Participants who had experienced notable growth in IC, as evidenced by the IDI scores, rated their own efforts to engage with the local context and population between 4 and 10, for an average of 6.6, while those showing more modest changes or negative growth rated these outreach efforts at 5.5 (ranging from 4 to 7). Two thirds of the participants who experienced positive gains on the IDI developmental orientation scores participated in extra-curricular activities offered at their host university. These activities were voluntary and distinct from the co-curricular program activities that are required for all participants and form an integral part of their capstone year. Participants chose to engage in a number of different activities such as discussion clubs, film club, English clubs with local students, music/singing, folk music instrument lessons and athletic events. Participants who developed close ties to host families and expanded their social network through attending events and activities gained more access to their host communities as well as more frequent exposure to local culture. Participants experiencing gains in IC reported significant contact with local individuals built through their networks of host families, other students at their host university and by introducing each other to those social group or groups that they were each able to join. These interactions occurred on a weekly basis, often taking place at least three times per week.

Conclusions

Several findings have emerged from this investigation, which are likely to be of interest to the field, as reflected in the research questions posed above.

(1) *How does OPI change differ across the immersion programs focused on early, mid and advanced levels of training and what, if any, relationships obtain between different levels of language proficiency and measured levels of intercultural development, as reflected by IDI?*

The present analysis has demonstrated that fixed-effect interactions between growth (f1) and program type are significant. While OPI change differences exist across the three programs, the average OPI change is approximately the same across all three programs. Language Flagship students attain growth comparable to that of lower- and middle-range-proficiency learners, but they produce these gains in the upper ranges of functional proficiency (ILR 3). The NSLI-Y group starts with the lowest

performance (of $7.80-4.46=3.34$) and demonstrates the largest improvement (of $1.83+1.63=3.46$). These interaction effects are several times larger than their standard errors and thus highly significant.

This study found no statistically significant relationship between IDI scores and OPI ratings for students at the ACTFL Novice, Intermediate and Advanced levels, but did observe a correlation between Superior-level OPI and higher IDI scores, limited somewhat by a small sample size at that level ($N=40$). It was also observed that mean post-program IDI scores throughout tended to be higher for those groups who began their overseas study with higher levels of speaking proficiency.

Evidence is presented here that early- and middle-stage students of language in the immersion context demonstrate significant competencies in their abilities to shift cognitive frames and to switch behavioral codes in response to local cultural contexts. If replicated elsewhere, this finding may provide a major rationale for early-stage immersion, as the finding held for both the language-specific groups (NSLI-Y and RLASP) and the non-language-oriented group (YES).

The OG, defined as the numerical difference between one's self-perception of intercultural development and actual measured developmental orientation, was observed to narrow as a result of immersion study at any level, with the narrowest gap between perception and actual level associated with those participants commanding the highest levels of language proficiency.

(2) *To what extent do pre-program IDI scores predict ultimate L2 attainment of SA participants?*

Pre-program IDI was *not* found to be a predictor of L2 proficiency gain at ACTFL Novice to Intermediate, Intermediate to Advanced or Advanced to Superior. While there is no doubt that the combination of attitudes and attributes measured by IDI are relevant for intercultural adjustment and adaptation in the overseas study context, the present analysis found no positive correlation of pre-program IDI scores and post-program OPI outcomes for any of the three levels or languages under study.

(3) *What is the effect of lower, mid-range and advanced L2 proficiency on change in IDI during SA?*

The study presents abundant data to support the position that overseas immersion with or without language study, but undertaken within a properly structured curriculum, contributes to intercultural development. It also provides evidence supporting the power of more advanced levels of language study to maintain and continue developmental growth over the period of a longer learning career, regardless of the learner's initial IDI level. While Cog and Beh are associated with higher developmental orientations, early-stage immersion has, nonetheless, been shown to produce significant growth in these areas for learners who have not yet attained the Acceptance or Adaptation orientations.

The data presented here strongly suggest that students at middle and higher levels of cultural learning and language competence become increasingly aware of their own behavior and communication in cross-cultural settings as they experience daily life and manage their interactions with local communities through program activities and guided reflection. Finally, the issues raised here also point to the need for further research on the optimization of the L2 speaker's intercultural self-awareness in connection with the strengthening of the critical comprehension, flexibility of mind and interactional skills on which all interculturality is based. Intercultural competency at the professional level is the centerpiece of the transversal skills so valued in today's globalized workforce and the *sine qua non* for a generation prepared to live responsibly in today's rapidly changing and highly interconnected world.

Notes

(1) The Assessment of Regional Knowledge (ARK) and the General Regional Aptitude Network Test (GRANT) test instruments currently under development at the US Military Academy, as referenced in Watson *et al.* (2013).

(2) The authors are pleased to acknowledge support in the modeling and review of these analyses by Dr Werner Wothke and Saodat Bazarova, American Councils Assessment Department.

(3) `>Model_Growth_times_Program_Interactions <- lmer(opi ~ 1 + (1 | id) + wave.f*prog.f)`

```
> display(Model_Growth_times_Program_Interactions)
```

```
lmer(formula = opi ~ 1 + (1 | id) + wave.f * prog.f)
```

```
coef.est coef.se
```

```
(Intercept) 7.80 0.21
```

```
wave.f1 1.83 0.20
```

```
prog.f2 -4.46 0.24
```

```
prog.f3 -2.91 0.24
```

```
wave.f1:prog.f2 1.63 0.23
```

```
wave.f1:prog.f3 -0.19 0.23
```

```
Error terms:
```

```
Groups Name Std.Dev.
```

```
id (Intercept) 0.95
```

```
Residual 0.89
```

```
---
```

```
number of obs: 514, groups: id, 275
```

```
AIC = 1672.8, DIC = 1630.9
```

```
deviance = 1643.9
```

Legend:		
Wave.f0:	Pre-program	assessment
Wave.f1:	Post-program	assessment
Prog.F1:		Flagship
Prog.F2:		NSLI-Y
Prog.F3:	RLASP	

References

- Anderson, P.H., Lawton, L., Rexeisen, R.J. and Hubbard, A.C. (2006) Short-term study abroad and intercultural sensitivity: A pilot study. *International Journal of Intercultural Relations* 30, 457–469.
- Anderson, P.H. and Lawton, L. (2011) Intercultural development: Study abroad vs. on-campus study. *Frontiers: The Interdisciplinary Journal of Study Abroad* 21, 86–108.
- Baker-Smemoe, D., Dewey, D.P., Brown, J. and Martinsen, R.A. (2014) Variables affecting L2 gains during study abroad. *Foreign Language Annals* 47 (3), 464–486.
- Bennett, M.J. (1986) A developmental approach to training for intercultural sensitivity. *International Journal of Intercultural Relations* 10 (2), 179–196.
- Bennett, M.J. (1993) Towards ethnorelativism: A developmental model of intercultural sensitivity. In R.M. Paige (ed.) *Education for the Intercultural Experience* (pp. 21–71). Yarmouth, ME: Intercultural Press.
- Bennett, M.J. (2004) Becoming interculturally competent. In J. Wurzel (ed.) *Toward Multiculturalism: A Reader in Multicultural Education* (2nd edn, pp. 62–77). Newton, MA: Intercultural Resource Corporation.
- Davidson, D.E. (2015) The development of L2 proficiency and literacy within the context of the federally supported overseas language training programs for Americans. In T. Brown and J. Bown (eds) *To Advanced Proficiency and Beyond: Theory and Methods for Developing Superior Second-Language Ability* (pp. 117–150). Washington, DC: Georgetown University Press.
- Davidson, D.E. and Lelic, M. (2012) The heritage and non-heritage learner in the overseas immersion context: Comparing learning outcomes and target-language utilization in the Russian flagship. *Russian Language Journal* 62, 47–77.
- Deardorff, D.K. (2004) The identification and assessment of intercultural competence as a student outcome of international education at institutions of higher education in the United States (Unpublished doctoral dissertation). North Carolina State University.
- Deardorff, D.K. (2006) Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education* 10 (3), 241–266.
- Engle, L. and Engle, J. (2004) Assessing language acquisition and intercultural sensitivity development in relation to study abroad program design. *Frontiers: The Interdisciplinary Journal of Study Abroad* 10, 219–236
- Freed, B.F., Segalowitz, N. and Dewey, D.P. (2004) Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition* 26 (2), 275–301.
- Hammer, M.R. (2007) *Saving Lives: The S.A.F.E. Model for Negotiating Hostage and Crisis Incidents*. Westport, CT: Praeger International Security.
- Hammer, M.R. (2012) The intercultural development inventory: A new frontier in assessment and development of intercultural competence. In M. Vande Berg, R.M. Paige and K.H. Lou (eds) *Student Learning Abroad* (pp. 115–136). Sterling, VA: Stylus Publishing.

- Hammer, M.R. (2015) *A Resource Guide for Effectively Using the Intercultural Development Inventory (IDI)*. Berlin, MD: IDI, LLC.
- Hammer, M.R., and Bennett, M.J. (1998) *The Intercultural Development Inventory (IDI) Manual*. Portland, OR: Intercultural Communication Institute.
- Hammer, M.R. and Bennett, M.J. (2002) *The Intercultural Development Inventory Manual*. Portland, OR: Intercultural Communication Institute.
- King, P.M. and Baxter Magolda, M. (2005) A developmental model of intercultural maturity. *Journal of College Student Development* 46, 571–592.
- Medina-Lopez-Portillo, A. (2004) Intercultural learning assessment: The link between program duration and the development of intercultural sensitivity. *Frontiers: The Interdisciplinary Journal of Study Abroad* 10, 179–200.
- Paige, R.M., Cohen, A.D. and Shively, R.L. (2004) Assessing the impact of a strategies-based curriculum on language and culture learning abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad* 10, 253–276.
- Patterson, P.K. (2006) Effect of study abroad on intercultural sensitivity (Unpublished doctoral dissertation). University of Missouri-Columbia.
- Pedersen, P.J. (2010) Assessing intercultural effectiveness outcomes in a year-long study abroad program. *International Journal of Intercultural Relations* 34 (1), 70–80.
- Rexeisen, R.J., Anderson, P., Lawton, L. and Hubbard, A. (2008) Study abroad and intercultural development: A longitudinal study. *Frontiers: The Interdisciplinary Journal of Study Abroad* 17, 1–20.
- Salisbury, M.H. (2011) The effect of study abroad on intercultural competence among undergraduate college students (Unpublished doctoral dissertation). University of Iowa.
- Vande Berg, M. (2009) Intervening in student learning abroad: A research-based inquiry. *Intercultural Education* 20 (1–2), 15–28.
- Vande Berg, M., Balkcum, A., Scheid, M. and Whalen, B. (2004) A report at the half-way mark: The Georgetown University consortium project. *Frontiers: The Interdisciplinary Journal of Study Abroad* 10, 101–116.
- Vande Berg, M., Connor-Linton, U.J. and Paige, M.R. (2009) The Georgetown consortium project: Interventions for student learning abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad* 18, 1–75.
- Watson, J.R., Siska, P. and Wolfel, R.L. (2013) Assessing gains in language proficiency, cross-cultural competence, and regional awareness during study abroad: A preliminary study. *Foreign Language Annals* 46 (1), 62–79.

Appendix 8.1: The Intercultural Development Continuum

This discussion of the IDI is adapted from Hammer (2015).

Denial: A Denial mindset reflects a more limited capability for understanding and appropriately responding to cultural differences in values, beliefs, perceptions, emotional responses and behaviors. Denial consists of a disinterest in other cultures and a more active avoidance of cultural difference. Individuals with a Denial orientation often do not see differences in perceptions and behavior as ‘cultural’. A Denial orientation is characteristic of individuals who have limited experience with other cultural groups and therefore tend to operate with broad stereotypes and generalizations about the cultural ‘other’. Those at Denial may also maintain a distance from other cultural groups and express little interest in learning about the cultural values and practices of diverse communities. This orientation tends to be associated more with members of a dominant culture as well as members of non-dominant groups who are relatively isolated from mainstream society, because both may have more opportunity to remain relatively isolated from cultural diversity. By contrast, members of non-dominant groups who are more actively engaged within the larger, mainstream society are less likely to maintain a Denial orientation, because they need to engage with cultural differences more often.

Polarization: Polarization is an evaluative mindset that views cultural differences from an ‘us versus them’ perspective. Polarization can take the form of defense (‘My cultural practices are superior to other cultural practices’) or reversal (‘Other cultures are better than mine’). Within defense, cultural differences are often seen as divisive and threatening to one’s own ‘way of doing things’. Reversal is a mindset that values and may idealize other cultural practices while denigrating one’s own culture group. Reversal may also support the ‘cause’ of an oppressed group, but this is done with little knowledge of what the ‘cause’ means to people from the oppressed community.

Minimization: Minimization is a transitional mindset between the more monocultural orientations of Denial and Polarization and the more intercultural/global worldviews of Acceptance and Adaptation. Minimization highlights commonalities in both human similarity (basic needs) and universalism (universal values and principles) that can mask a deeper understanding of cultural differences. Minimization can take one of two forms: (a) the highlighting of commonalities due to limited cultural self-understanding, which is more commonly experienced by dominant group members within a cultural community; or (b) the highlighting of commonalities as a strategy for navigating the values and practices largely determined by the dominant culture group, which is more often experienced

by non-dominant group members within a larger cultural community. This latter strategy can have survival value for non-dominant culture members and often takes the form of 'go along to get along'.

Acceptance: Acceptance and Adaptation are intercultural/global mindsets. With an Acceptance orientation, individuals recognize and appreciate patterns of cultural difference and commonality in their own and other cultures. An Acceptance orientation is curious to learn how a cultural pattern of behavior makes sense within different cultural communities. This involves contrastive self-reflection between one's own culturally learned perceptions and behaviors and perceptions and practices of different cultural groups. While curious, individuals with an Acceptance mindset are not fully able to appropriately adapt to cultural difference. Someone with an Acceptance orientation may be challenged as well to make ethical or moral decisions across cultural groups. While a person within Acceptance embraces a deeper understanding of cultural differences, this can lead to the individual struggling with reconciling behavior in another cultural group that the person considers unethical or immoral from his or her own cultural viewpoint.

Adaptation: An Adaptation orientation consists of both Cog (shifting one's cultural perspective) and Beh (changing behavior in authentic and culturally appropriate ways). Adaptation enables deep cultural bridging across diverse communities using an increased repertoire of cultural frameworks and practices in navigating cultural commonalities and differences. An Adaptation mindset sees adaptation in performance (behavior). While people with an Adaptation mindset typically focus on learning adaptive strategies, problems can arise when people with Adaptation mindsets express little tolerance toward people who engage diversity from other developmental orientations. This can result in people with Adaptive capabilities being marginalized in their workplace.